



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/815,354	03/31/2004	Patrick Chiu	FXPL-01094USO	8294
23910 7590 08/24/2009 FLIESLER MEYER LLP 650 CALIFORNIA STREET 14TH FLOOR SAN FRANCISCO, CA 94108				
EXAMINER PARK, EDWARD				
ART UNIT 2624		PAPER NUMBER		
NOTIFICATION DATE 08/24/2009		DELIVERY MODE ELECTRONIC		

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

OFFICEACTIONS@FDML.COM

Office Action Summary

Application No.

10/815,354

Applicant(s)

CHIU ET AL.

Examiner

EDWARD PARK

Art Unit

2624

Period for Reply -- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 15 May 2009.
- 2a) ☒ This action is **FINAL**. 2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-6, 8-15 and 17-27 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1-6, 8-15, 17-27 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
 2. ☐ Certified copies of the priority documents have been received in Application No. _____.
 3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☐ Information Disclosure Statement(s) (PTO-8508)
Paper No(s)/Mail Date _____
- 4) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date _____
- 5) ☐ Notice of Informal Patent Application
- 6) ☐ Other: _____

DETAILED ACTION

Response to Amendment

1. This action is responsive to applicant's amendment and remarks received on 5/15/09. Claims 1-6, 8-15, 17-27 are currently pending.

Response to Arguments

2. Applicant's arguments with respect to claim 1 have been considered but are moot in view of the new ground(s) of rejection. Applicant argues that the Bae reference does not disclose the cited limitations within claim 1 (see pg. 12, first paragraph - pg. 19, second paragraph). This argument is considered moot in view of a new ground(s) of rejection which is necessitated by applicant's amendment of claim 1, and the rejection can be seen below.

Applicant's arguments with respect to claim 10 have been considered but are moot in view of the new ground(s) of rejection. Applicant argues that the Bae reference does not disclose the cited limitations within claim 10 (see pg. 19, third paragraph - pg. 24, first paragraph). This argument is considered moot in view of a new ground(s) of rejection which is necessitated by applicant's amendment of claim 10, and the rejection can be seen below.

Regarding claim 12, applicant argues that the claim is allowable due to the same reasons as stated within claim 10 (see pg. 12, second paragraph). This argument is not considered persuasive since claim 10 stands rejected under a new ground(s) of rejection necessitated by applicant's amendment and the rejection can be seen below.

Regarding claim 22, applicant argues that Bae does not define a support (see pg. 24, last paragraph). This argument is not considered persuasive since Bae discloses this limitation within fig. 5, paragraph [0059]; summarizing important texts in a video frame in each of the news articles. Examiner notes that the germ is considered to be the text and the support is the background surrounding the text. Applicant argues that Bae does not implicitly define a support in each of the video segments (see pg. 25, first paragraph – third paragraph). This argument is not considered persuasive since it is seen within figure 5, 6 paragraphs [0059], [0051], where synthetic text key frame is used to provide a summary of important video text included in each of the articles. It is shown that the region of interest/germ is the text and the support is the background behind the text. Applicant argues that there is no explicit disclosure of separating the germ from the video segments in Bae (see pg. 26, first paragraph). This argument is not considered persuasive since Bae discloses separating the germs within fig. 5, paragraph [0059], where the synthetic text key frame is generated by summarizing important texts in a video frame in each of the news articles to be inverted into an image. It is seen within figure 5 that the key frame there is an extraction of the germ including the support to synthesize a synthetic text key frame. Applicant argues that Bae does not implicitly define separating the germ from the video segments (see pg. 26, second paragraph - pg. 27, second paragraph). This argument is not considered persuasive since the previous argument addresses this limitation and the applicant is directed to the response as seen above.

Applicant's arguments with respect to claim 23 have been considered but are moot in view of the new ground(s) of rejection. Applicant argues that the Bae reference does not disclose the cited limitations within claim 23 (see pg. 27, third paragraph - pg. 32, second

paragraph). This argument is considered moot in view of a new ground(s) of rejection which is necessitated by applicant's amendment of claim 23, and the rejection can be seen below.

Applicant's arguments with respect to claim 24 have been considered but are moot in view of the new ground(s) of rejection. Applicant argues that the Bae reference does not disclose the cited limitations within claim 24 (see pg. 32, third paragraph - pg. 32, last paragraph). This argument is considered moot in view of a new ground(s) of rejection which is necessitated by applicant's amendment of claim 24, and the rejection can be seen below.

Regarding claim 2, applicant argues that Bae does needs to be modified to determine a dominant group that has the largest 3-D volume (see pg. 33, last paragraph). This argument is not considered persuasive since the rejection of claim 2 addresses this limitation as seen below in the rejection. Applicant argues that claim 2 is directed to a volume, not time duration and therefore Uchihashi does not teach the limitation (see pg. 34, second paragraph). This argument is not considered persuasive since the term 3-D volume can be interpreted as reasonable broad as possible and comprises an x, y, t dimension that would constitutes a certain quantity that has three dimensions.

Regarding claims 3-6, 13-15, 20, applicant argues that the claims are allowable due to the same reasons as stated within claims 1, 10 (see pg. 34, third paragraph). This argument is not considered persuasive since claims 1, 10 stand rejected under a new ground(s) of rejection necessitated by applicant's amendment as seen within this action.

Applicant's arguments with respect to claims 7, 9, 16, 18 have been considered but are moot in view of the new ground(s) of rejection. Applicant argues that the Hirata reference does not disclose the cited limitations within claims 9, 18 (see pg. 34, last paragraph - pg. 37, second

paragraph). This argument is considered moot in view of a new ground(s) of rejection which is necessitated by applicant's amendment of independent claims 1, 10, and the rejections can be seen below.

Regarding claims 11, 19, applicant argues that the claims are allowable due to the same reasons as stated within claims 1 and 10 (see pg. 37, fourth paragraph). This argument is not considered persuasive since claims 1, 10 stand rejected under a new ground(s) of rejection necessitated by applicant's amendment as seen within this action.

Regarding claim 21, applicant argues that it would not have been obvious for one of ordinary skill in the art to adapt the teaching of 3-D coordinates to two dimensional pictures (see pg. 37, last paragraph—pg. 38, last paragraph). This argument is not considered persuasive since Leow only brings in the concept of computing boundary curves between the germs using a Voronoi algorithm to be incorporated with Bae, no more or less. In response to applicant's argument that it would not have been obvious to combine Leow with Bae, the test for obviousness is not whether the features of a secondary reference may be bodily incorporated into the structure of the primary reference; nor is it that the claimed invention must be expressly suggested in any one or all of the references. Rather, the test is what the combined teachings of the references would have suggested to those of ordinary skill in the art. See *In re Keller*, 642 F.2d 413, 208 USPQ 871 (CCPA 1981). In response to applicant's argument that there is no suggestion to combine the references, the examiner recognizes that obviousness can only be established by combining or modifying the teachings of the prior art to produce the claimed invention where there is some teaching, suggestion, or motivation to do so found either in the references themselves or in the knowledge generally available to one of ordinary skill in the art.

See *In re Fine*, 837 F.2d 1071, 5 USPQ2d 1596 (Fed. Cir. 1988) and *In re Jones*, 958 F.2d 347, 21 USPQ2d 1941 (Fed. Cir. 1992). In this case, it would have been obvious to one ordinary skilled in the art to modify the Bae reference to utilize a Voronoi algorithm to compute boundary curves as taught by Leow, in order for the output to be topologically correct and convergent to the original surface as the sampling density increases for a "good sample" from a smooth surface (see col. 2, lines 19-50).

Regarding claim 25, applicant argues that it would not have been obvious for one of ordinary skill in the art to adapt the teaching of 3-D coordinates to two dimensional pictures (see pg. 39, last paragraph—pg. 40, second paragraph). This argument is not considered persuasive since Leow only brings in the concept of computing boundary curves between the germs using a Voronoi algorithm to be incorporated with Bae, no more or less. In response to applicant's argument that it would not have been obvious to combine Leow with Bae, the test for obviousness is not whether the features of a secondary reference may be bodily incorporated into the structure of the primary reference; nor is it that the claimed invention must be expressly suggested in any one or all of the references. Rather, the test is what the combined teachings of the references would have suggested to those of ordinary skill in the art. See *In re Keller*, 642 F.2d 413, 208 USPQ 871 (CCPA 1981). In response to applicant's argument that there is no suggestion to combine the references, the examiner recognizes that obviousness can only be established by combining or modifying the teachings of the prior art to produce the claimed invention where there is some teaching, suggestion, or motivation to do so found either in the references themselves or in the knowledge generally available to one of ordinary skill in the art. See *In re Fine*, 837 F.2d 1071, 5 USPQ2d 1596 (Fed. Cir. 1988) and *In re Jones*, 958 F.2d 347,

21 USPQ2d 1941 (Fed. Cir. 1992). In this case, it would have been obvious to one ordinary skilled in the art to modify the Bae reference to utilize a Voronoi algorithm to compute boundary curves as taught by Leow, in order for the output to be topologically correct and convergent to the original surface as the sampling density increases for a "good sample" from a smooth surface (see col. 2, lines 19-50).

Claim Rejections - 35 USC § 101

3. In response to applicant's amendment of claims 1-25, the previous claim rejection is withdrawn.

Claim Rejections - 35 USC § 102

4. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(a) the invention was known or used by others in this country, or patented or described in a printed publication in this or a foreign country, before the invention thereof by the applicant for a patent.

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

(c) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

5. **Claim 22** is rejected under 35 U.S.C. 102(b) as being anticipated by Bae et al (US 2002/0126143 A1).

Regarding **claim 22**, Bac discloses a computer implemented method implemented within a computer system including memory and CPU for generating a highly condensed visual summary of video regions, comprising:

utilizing the memory and CPU for determining a germ in each of a plurality of images, the germ containing a region of interest (see fig. 5, paragraph [0059]; The synthetic text key frame is generated by summarizing important texts in a video frame in each of the news articles to be inverted into an image; Examiner notes within figure 5, it is noted that the physical text is considered as the germ);

utilizing the memory and CPU for defining a support in each of the video segments, wherein the support is the video segment less the germ (see fig. 5, paragraph [0059]; summarizing important texts in a video frame in each of the news articles. Examiner notes that the germ is considered to be the text and the support is the background surrounding the text);

utilizing the memory and CPU for separating the germ from the video segments (see fig. 5, paragraph [0059]; The synthetic text key frame is generated by summarizing important texts in a video frame in each of the news articles to be inverted into an image);

utilizing the memory and CPU for laying out the germs on a canvas (see fig. 5, paragraph [0059]; the synthetic text key frame is generated by summarizing important texts), wherein there is no more than one germ for every video segment; and utilizing the memory and CPU for filling in the space of the canvas between the germs (see fig. 5, paragraph [0059]; The synthetic text key frame is generated by summarizing important texts in a video frame in each of the news articles to be inverted into an image; Examiner notes that as seen in fig. 5, the germ/region of interest is interpreted to be the text only. Therefore, the corresponding background surrounding the text is

considered to be support, since the support is defined as the video segment less the germ.

Following this logic, therefore when the summarization of the text occurs, the support along with the germ is separated from the video segments and placed within the synthetic key frame, whereby it fills up the space of the canvas between the germs by having more than one text summarization present in the synthetic key frame that represents only one germ for every video segment/anchor frame/icon key frame. Inherently, if the part of the support and the germ are transferred to the synthetic key frame then naturally, at least one pixel value of the support relative to the closes germ is positioned corresponding to the position of that pixel value relative to the germ) to generate a highly condensed visual summary of the plurality of video segments (see fig. 5, paragraph [0059], synthetic key frame).

6. **Claim 23** are rejected under 35 U.S.C. 102(b) as being anticipated by Yu et al (US 2002/0126203 A1).

Regarding **claim 23**, Yu discloses a computer implemented method implemented within a computer system including memory and CPU for generating a highly condensed visual summary of video regions, comprising:

utilizing the memory and CPU for determining a germ in each of a plurality of images, the germ containing a region of interest (see paragraph [0013]; most of video indexing systems extract key frames to represent the scenes and shots as the structural components of the video stream, and use the same for the purpose of searching or browsing. In order to efficiently carry out the foregoing process);

utilizing the memory and CPU for defining a support in each of the video segments, wherein the support is the video segment less the germ (see fig. 3, paragraph [0047]; text area is extracted, a

weight is determined to the extracted text area (step 13). The weight is determined by using weight determining factors, which may include the size of the text area, the mean text size in the text area, the display duration time of a text and the like. Therefore, the weight can be determined in proportion to the size of the text area, the mean text size in the text area and the display duration time of the text. In other words, as the size of the text area or the mean text size in the text area increases, the weight can increase also. In the same manner, as the display duration time increases, the weight can increase. Of course, when each weight determining factor decreases or reduces, the weight can proportionally decrease);

utilizing the memory and CPU for separating the germ from the video segments (see fig. 4, paragraph [0050]; a synthetic key frame can be generated by synthesizing only a preferred text area among the text areas extracted from the video stream with the key frame according to an importance measure satisfying an importance function);

utilizing the memory and CPU for laying out the germs on a canvas, wherein the germs are laid out in irregular two dimensional shapes on the canvas (see fig. 4, paragraph [0054]; importance of the text area is compared with a pre-set importance (step 17). The pre-set importance can be set according to the size of a device to be displayed or the size of the synthetic key frame area in a browser. If the size of the browser increases, the size of the synthetic key frame can be increased. Accordingly, the number or size of the text areas to be synthesized can be increased and the importance measure can be also increased. If the number or size of the key frame to be synthesized is changed);

utilizing the memory and CPU for defining a space between the germs (see paragraph [0054]; importance of the text area is compared with a pre-set importance (step 17). The pre-set

importance can be set according to the size of a device to be displayed or the size of the synthetic key frame area in a browser. If the size of the browser increases, the size of the synthetic key frame can be increased. Accordingly, the number or size of the text areas to be synthesized can be increased and the importance measure can be also increased. If the number or size of the key frame to be synthesized is changed, the readability of the user can be considered); and filling in the space of the canvas between the germs, wherein filling in the space of the canvas between the germs includes laying out one or more portions of the supports (see fig. 4, paragraph [0058]; the synthetic key frame generated in step 21 is generated for the text areas extracted from one shot or scene, so that the steps 11 to 21 are repeatedly performed to generate one synthetic key frame per one shot or scene included in the video stream), wherein the canvas generated is a highly condensed visual summary of the plurality of video segments (see fig. 5, paragraph [0059], generating a synthetic key frame based upon video text about a specific article interval in a news video, and FIG. 6 illustrates a method of generating a synthetic key frame based upon video text in a show program).

Claim Rejections - 35 USC § 103

7. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

8. **Claims 1, 8, 10, 12, 17** are rejected under 35 U.S.C. 103(a) as being unpatentable over Yu et al (US 2002/0126203 A1) in view of Yasui (US 6,081,615).

Regarding **claim 1**, Yu discloses a computer implemented method implemented within a computer system including memory and CPU for generating a highly condensed visual summary of video regions, comprising:

utilizing the memory and CPU for determining a dominant group in each of a plurality of video segments (see paragraph [0010], [0013]; conventional video indexing techniques structurally analyze the video stream to detect the shots and scenes as unit segments and extract key frames based upon the shots and scenes. The key frames represent the shots and scenes, and those key frames are utilized as a material for summarizing the video or used as means for moving to desired positions; most of video indexing systems extract key frames to represent the scenes and shots as the structural components of the video stream, and use the same for the purpose of searching or browsing);

utilizing the memory and CPU for determining a key frame in each of the video segments (see paragraph [0013]; most of video indexing systems extract key frames to represent the scenes and shots as the structural components of the video stream, and use the same for the purpose of searching or browsing. In order to efficiently carry out the foregoing process);

utilizing the memory and CPU for defining a germ associated with each dominant group in each of the video segments, wherein the video segment less the germ defines a support in each of the video segments (see fig. 3, paragraph [0047]; text area is extracted, a weight is determined to the extracted text area (step 13). The weight is determined by using weight determining factors, which may include the size of the text area, the mean text size in the text area, the display

duration time of a text and the like. Therefore, the weight can be determined in proportion to the size of the text area, the mean text size in the text area and the display duration time of the text. In other words, as the size of the text area or the mean text size in the text area increases, the weight can increase also. In the same manner, as the display duration time increases, the weight can increase. Of course, when each weight determining factor decreases or reduces, the weight can proportionally decrease);

utilizing the memory and CPU for separating the germ from the video segments (see fig. 4, paragraph [0050]; a synthetic key frame can be generated by synthesizing only a preferred text area among the text areas extracted from the video stream with the key frame according to an importance measure satisfying an importance function);

utilizing the memory and CPU for laying out the germs on a canvas (see fig. 4, paragraph [0054]; importance of the text area is compared with a pre-set importance (step 17). The pre-set importance can be set according to the size of a device to be displayed or the size of the synthetic key frame area in a browser. If the size of the browser increases, the size of the synthetic key frame can be increased. Accordingly, the number or size of the text areas to be synthesized can be increased and the importance measure can be also increased. If the number or size of the key frame to be synthesized is changed); and

utilizing the memory and CPU for filling in the space of the canvas between the germs, wherein filling in the space of the canvas between the germs includes laying out one or more portions of the supports (see fig. 4, paragraph [0058]; the synthetic key frame generated in step 21 is generated for the text areas extracted from one shot or scene, so that the steps 11 to 21 are repeatedly performed to generate one synthetic key frame per one shot or scene included in the

video stream), wherein the canvas generated is a highly condensed visual summary of the plurality of video segments (see fig. 5, paragraph [0059], generating a synthetic key frame based upon video text about a specific article interval in a news video, and FIG. 6 illustrates a method of generating a synthetic key frame based upon video text in a show program). Yu does not disclose assigning a pixel value of a point in the space from pixel values of a support of a neighboring germ based on a distance from the point to the neighboring germ.

Yasui, in the same field of endeavor, teaches assigning a pixel value of a point in the space from pixel values of a support of a neighboring germ based on a distance from the point to the neighboring germ (see fig. 12, col. 3, lines 34-55; interpolation whereby during mapping of texture cells representing the surface pattern of the object which is to be displayed, colour boundaries are interpolated by blending the colours of the texture cells contained within prescribed limits, wherein the process of interpolation comprises the steps of: calculating the magnification ratio of the texture cells in relation to the pixels; determining the size of filter for stipulating the area which is subject to interpolation in accordance with the calculated magnification ratio; determining position, whereby a filter of the stipulated size is moved one by one over the texture cells which are to be mapped, and the position is determined; extraction, whereby the colour values of the texture cells which overlap with the filter in each position are extracted; blending colour values, whereby extracted colour values for every texture cell are blended in proportion to the area of overlap between the texture cell in question and the filter; and mapping, whereby the blended colour values are mapped on each pixel which corresponds to a position of the filter).

It would have been obvious at the time the invention was made to one of ordinary skill in the art to modify the Yu reference to assign a pixel value based on distance from a germ as suggested by Yasui, to prevent excessive fuzziness when colour boundaries are interpolated, and to allow color changes in the edge sections to become smoother and jaggy appearances are reduced (see col. 1, lines 6-12, col. 2, lines 16-23).

Regarding **claim 8**, Yu with Yasui discloses all the elements as mentioned above in claim 1. Yu with Yasui does not teach the point is assigned the pixel value of the closest germ with a support that includes the point.

Yasui, in the same field of endeavor, teaches the point is assigned the pixel value of the closest germ with a support that includes the point (see fig. 12, col. 3, lines 10-55).

It would have been obvious at the time the invention was made to one of ordinary skill in the art to modify the Yu with Yasui to assign a pixel value based closest to a germ as suggested by Yasui, to prevent excessive fuzziness when colour boundaries are interpolated, and to allow color changes in the edge sections to become smoother and jaggy appearances are reduced (see col. 1, lines 6-12, col. 2, lines 16-23).

Regarding **claim 10**, Yu discloses a computer implemented method implemented within a computer system including memory and CPU for generating a highly condensed visual summary of video regions, comprising:
utilizing the memory and CPU for determining a germ in each of a plurality of images, the germ containing a region of interest, wherein the video region less the germ defines a support in each of the video regions (see fig. 3, paragraph [0047]; text area is extracted, a weight is determined to the extracted text area (step 13). The weight is determined by using weight determining factors,

which may include the size of the text area, the mean text size in the text area, the display duration time of a text and the like. Therefore, the weight can be determined in proportion to the size of the text area, the mean text size in the text area and the display duration time of the text. In other words, as the size of the text area or the mean text size in the text area increases, the weight can increase also. In the same manner, as the display duration time increases, the weight can increase. Of course, when each weight determining factor decreases or reduces, the weight can proportionally decrease);

utilizing the memory and CPU for separating the germ from the video segments (see fig. 4, paragraph [0050]; a synthetic key frame can be generated by synthesizing only a preferred text area among the text areas extracted from the video stream with the key frame according to an importance measure satisfying an importance function);

utilizing the memory and CPU for laying out the germs on a canvas, wherein the germs are laid out in irregular two dimensional shapes on the canvas (see fig. 4, paragraph [0054]; importance of the text area is compared with a pre-set importance (step 17). The pre-set importance can be set according to the size of a device to be displayed or the size of the synthetic key frame area in a browser. If the size of the browser increases, the size of the synthetic key frame can be increased. Accordingly, the number or size of the text areas to be synthesized can be increased and the importance measure can be also increased. If the number or size of the key frame to be synthesized is changed); and

utilizing the memory and CPU for filling in the space of the canvas between the irregular two dimensional shape germs by laying out one or more parts of the support (see fig. 4, paragraph [0058]; the synthetic key frame generated in step 21 is generated for the text areas extracted from

one shot or scene, so that the steps 11 to 21 are repeatedly performed to generate one synthetic key frame per one shot or scene included in the video stream), wherein the canvas generated is a highly condensed visual summary of video regions (see fig. 5, paragraph [0059], generating a synthetic key frame based upon video text about a specific article interval in a news video, and FIG. 6 illustrates a method of generating a synthetic key frame based upon video text in a show program). Yu does not disclose assigning a pixel value of a point in the space from pixel values of a support of a neighboring germ based on a distance from the point to the neighboring germ.

Yasui, in the same field of endeavor, teaches assigning a pixel value of a point in the space from pixel values of a support of a neighboring germ based on a distance from the point to the neighboring germ (see fig. 12, col. 3, lines 34-55; interpolation whereby during mapping of texture cells representing the surface pattern of the object which is to be displayed, colour boundaries are interpolated by blending the colours of the texture cells contained within prescribed limits, wherein the process of interpolation comprises the steps of: calculating the magnification ratio of the texture cells in relation to the pixels; determining the size of filter for stipulating the area which is subject to interpolation in accordance with the calculated magnification ratio; determining position, whereby a filter of the stipulated size is moved one by one over the texture cells which are to be mapped, and the position is determined; extraction, whereby the colour values of the texture cells which overlap with the filter in each position are extracted; blending colour values, whereby extracted colour values for every texture cell are blended in proportion to the area of overlap between the texture cell in question and the filter; and mapping, whereby the blended colour values are mapped on each pixel which corresponds to a position of the filter).

It would have been obvious at the time the invention was made to one of ordinary skill in the art to modify the Yu reference to assign a pixel value based on distance from a germ as suggested by Yasui, to prevent excessive fuzziness when colour boundaries are interpolated, and to allow color changes in the edge sections to become smoother and jaggy appearances are reduced (see col. 1, lines 6-12, col. 2, lines 16-23).

Regarding **claim 12**, Yu discloses receiving user input, the user input associated with a part of an image (see paragraph [0065]).

Regarding **claim 17**, Yu with Yasui discloses all the elements as mentioned above in claim 10. Yu with Yasui does not teach the point is assigned the pixel value of the closest germ with a support that includes the point.

Yasui, in the same field of endeavor, teaches the point is assigned the pixel value of the closest germ with a support that includes the point (see fig. 12, col. 3, lines 10-55).

It would have been obvious at the time the invention was made to one of ordinary skill in the art to modify the Yu with Yasui to assign a pixel value based closest to a germ as suggested by Yasui, to prevent excessive fuzziness when colour boundaries are interpolated, and to allow color changes in the edge sections to become smoother and jaggy appearances are reduced (see col. 1, lines 6-12, col. 2, lines 16-23).

9. **Claim 24** is rejected under 35 U.S.C. 103(a) as being unpatentable over Yu et al (US 2002/0126203 A1) in view of Li et al (US 7,035,435 B2).

Regarding **claim 24**, Yu discloses a computer implemented method implemented within a computer system including memory and CPU for generating a highly condensed visual summary of video regions, comprising:

utilizing the memory and CPU for determining a dominant group in each of a plurality of video segments (see paragraph [0010], [0013]; conventional video indexing techniques structurally analyze the video stream to detect the shots and scenes as unit segments and extract key frames based upon the shots and scenes. The key frames represent the shots and scenes, and those key frames are utilized as a material for summarizing the video or used as means for moving to desired positions; most of video indexing systems extract key frames to represent the scenes and shots as the structural components of the video stream, and use the same for the purpose of searching or browsing);

utilizing the memory and CPU for determining a key frame in each of the video segments (see paragraph [0013]; most of video indexing systems extract key frames to represent the scenes and shots as the structural components of the video stream, and use the same for the purpose of searching or browsing. In order to efficiently carry out the foregoing process);

utilizing the memory and CPU for defining a germ associated with each dominant group in each of the video segments, wherein the germ is the x-y projection of the dominant group onto the keyframe (see fig. 3, paragraph [0047]; text area is extracted, a weight is determined to the extracted text area (step 13). The weight is determined by using weight determining factors, which may include the size of the text area, the mean text size in the text area, the display duration time of a text and the like. Therefore, the weight can be determined in proportion to the size of the text area, the mean text size in the text area and the display duration time of the text. In other words, as the size of the text area or the mean text size in the text area increases, the weight can increase also. In the same manner, as the display duration time increases, the weight can

increase. Of course, when each weight determining factor decreases or reduces, the weight can proportionally decrease);

utilizing the memory and CPU for separating the germ from the video segments (see fig. 4, paragraph [0050]; a synthetic key frame can be generated by synthesizing only a preferred text area among the text areas extracted from the video stream with the key frame according to an importance measure satisfying an importance function);

utilizing the memory and CPU for laying out the germs on a canvas (see fig. 4, paragraph [0054]; importance of the text area is compared with a pre-set importance (step 17). The pre-set importance can be set according to the size of a device to be displayed or the size of the synthetic key frame area in a browser. If the size of the browser increases, the size of the synthetic key frame can be increased. Accordingly, the number or size of the text areas to be synthesized can be increased and the importance measure can be also increased. If the number or size of the key frame to be synthesized is changed); and

utilizing the memory and CPU for filling in the space of the canvas between the germs (see fig. 4, paragraph [0058]; the synthetic key frame generated in step 21 is generated for the text areas extracted from one shot or scene, so that the steps 11 to 21 are repeatedly performed to generate one synthetic key frame per one shot or scene included in the video stream), wherein the canvas generated is a highly condensed visual summary of the plurality of video segments (see fig. 5, paragraph [0059], generating a synthetic key frame based upon video text about a specific article interval in a news video, and FIG. 6 illustrates a method of generating a synthetic key frame based upon video text in a show program). Yu does not disclose dominant group that includes a face.

Li teaches dominant group that includes a face (Li: col. 7, lines 33-51).

It would have been obvious at the time the invention was made to one of ordinary skill in the art to modify the Yu reference to detect a face as taught by Li, in order to determine the importance of a frame since “a human face will be more informative than, for example, a landscape frame” (Li: col. 7, lines 33-51).

10. **Claims 2-6, 13-15, 20, 26** are rejected under 35 U.S.C. 103(a) as being unpatentable over Yu et al (US 2002/0126203 A1) with Yasui (US 6,081,615), and further in view of Uchihashi (ACM Multimedia: “Video Manga: Generating Semantically Meaningful Video Summaries”).

Regarding **claim 2**, Yu with Yasui discloses all elements as mentioned above in claim 1. Yu with Yasui does not teach determining a group within each of the plurality of video segments having the largest volume.

Uchihashi teaches determining a group within each of the plurality of video segments having the largest 3-D volume (Uchihashi: section 4.2, length of the segment is scored).

It would have been obvious at the time the invention was made to one of ordinary skill in the art to modify Yu with Yasui to determine a group having the largest volume as taught by Uchihashi, in order to “calculate an importance score for each segment based on its rarity and duration” since “a segment is deemed less important if it is short or very similar to other segments” (Uchihashi: section 4.2).

Regarding **claims 3, 4, and 20**, Yu with Yasui discloses all elements as mentioned above in claim 1. Yu with Yasui does not teach defining a two dimensional shape that encompasses the projection of the dominant group onto the key frame; wherein the two dimensional shape is a rectangle; and using an algorithm to determine a region of interest of an image.

Uchihashi teaches defining a two dimensional shape that encompasses the projection of the dominant group onto the key frame (Uchihashi: figure 2; section 4.4) and wherein the two dimensional shape is a rectangle (Uchihashi: figure 2; section 4.4).

It would have been obvious at the time the invention was made to one of ordinary skill in the art to modify Yu with Yasui to define a two dimensional shape that is a rectangle as taught by Uchihashi, in order to “form a pictorial abstract of the video sequence” where a “sequence of frames fills space efficiently and represents the original video sequence well” (Uchihashi: section 4.4).

Uchihashi further teaches using an algorithm to determine a region of interest of an image (Uchihashi: figure 4.2).

It would have been obvious at the time the invention was made to one of ordinary skill in the art to modify the Yu with Yasui with Uchihashi combination as mentioned above to determine a region of interest of an image as taught by Uchihashi, “to select appropriate keyframes for a compact pictorial summary” (Uchihashi: section 4.2).

Regarding **claims 5 and 6**, Yu with Yasui with Uchihashi discloses all elements as mentioned above in claim 3. Yu with Yasui with Uchihashi as mentioned in claim 3, does not teach determining a scale factor to be applied to every germ such that the germs are scaled to the maximum size that fits into the canvas and placing the germs in rows, wherein each row has a height according to the longest germ in the particular row.

Uchihashi further teaches determining a scale factor to be applied to every germ such that the germs are scaled to the maximum size that fits into the canvas (Uchihashi: section 4.3, 4.4)

and placing the germs in rows, wherein each row has a height according to the longest germ in the particular row (Uchihashi: figure 2).

It would have been obvious at the time the invention was made to one of ordinary skill in the art to modify the Yu with Yasui with Uchihashi combination to place the germs in a row as taught by Uchihashi, to “fill space efficiently and represent the original video sequence well” (Uchihashi: section 4.2).

Regarding **claim 13**, Yu with Yasui discloses all elements as mentioned above in claim 10. Yu with Yasui does not disclose using an algorithm to determine the regions of interest of an image based on one or more methods selected from the group consisting of a general image analysis algorithm, a face-detection algorithm, and object detection algorithms and user input.

Uchihashi teaches using an algorithm to determine the regions of interest of an image based on one or more methods selected from the group consisting of a face-detection algorithm (see section 4.2, section 6, segment is scored and weighted, detecting human close-ups and other image types to further improve the summaries), and object detection algorithms (see section 4.2, section 6, segment is scored and weighted, detecting human close-ups and other image types to further improve the summaries) and user input.

It would have been obvious at the time the invention was made to one of ordinary skill in the art to modify Yu with Yasui to determine the regions of interest with a face-detection algorithm or a object detection algorithms as taught by Uchihashi, “to select appropriate keyframes for a compact pictorial summary” (Uchihashi: section 4.2).

Regarding **claims 14 and 15**, Yu with Yasui reference discloses all elements as mentioned above in claim 10. Yu with Yasui reference as mentioned in claim 10, does not teach

determining a scale factor to be applied to every germ such that the germs are scaled to the maximum size that fits into the canvas and placing the germs in rows, wherein each row has a height according to the longest germ in the particular row.

Uchihashi further teaches determining a scale factor to be applied to every germ such that the germs are scaled to the maximum size that fits into the canvas (Uchihashi: section 4.3, 4.4) and placing the germs in rows, wherein each row has a height according to the longest germ in the particular row (Uchihashi: figure 2).

It would have been obvious at the time the invention was made to one of ordinary skill in the art to modify Yu with Yasui to place the germs in a row as taught by Uchihashi, to “fill space efficiently and represent the original video sequence well” (Uchihashi: section 4.2).

Regarding **claim 26**, Yu further discloses two dimensional shape is irregular (see figs. 4-6).

11. **Claims 9, 18** are rejected under 35 U.S.C. 103(a) as being unpatentable over Yu et al (US 2002/0126203 A1) with Yasui (US 6,081,615), and further in view of Kasamatsu (US 5,761,338).

Regarding **claim 9**, Yu with Yasui discloses all elements as mentioned above in claim 1. Yu with Yasui does not teach wherein the point is assigned a background value if no support includes the point.

Hirata teaches wherein the point is assigned a background value if no support includes the point (see col. 9, lines 10-34).

It would have been obvious at the time the invention was made to one of ordinary skill in the art to modify the Yu with Yasui to assign a background value as taught by Kasamatsu, to

accurately detect an image area of a document based on the density of the background portion of the document (see col. 1, lines 35-58).

Regarding **claim 18**, Yu with Yasui discloses all elements as mentioned above in claim 10. Yu with Yasui does not teach wherein the point is assigned a background value if no support includes the point.

Hirata teaches wherein the point is assigned a background value if no support includes the point (see col. 9, lines 10-34).

It would have been obvious at the time the invention was made to one of ordinary skill in the art to modify the Yu with Yasui to assign a background value as taught by Kasamatsu, to accurately detect an image area of a document based on the density of the background portion of the document (see col. 1, lines 35-58).

12. **Claims 11, 19** is rejected under 35 U.S.C. 103(a) as being unpatentable over Yu et al (US 2002/0126203 A1) with Yasui (US 6,081,615), and further in view of Li et al (US 7,035,435 B2).

Regarding **claim 11**, Yu with Yasui discloses all elements as mentioned above in claim 10. Yu with Yasui does not teach detecting a face in each of the plurality of images.

Li teaches detecting a face in each of the plurality of images (Li: col. 7, lines 33-51).

It would have been obvious at the time the invention was made to one of ordinary skill in the art to modify Yu with Yasui to detect a face as taught by Li, in order to determine the importance of a frame since “a human face will be more informative than, for example, a landscape frame” (Li: col. 7, lines 33-51).

Regarding **claim 19**, Yu with Yasui discloses all elements as mentioned above in claim 1. Yu with Yasui does not teach detecting a face in each of the plurality of images.

Li teaches detecting a face in each of the plurality of images (Li: col. 7, lines 33-51).

It would have been obvious at the time the invention was made to one of ordinary skill in the art to modify Yu with Yasui to detect a face as taught by Li, in order to determine the importance of a frame since “a human face will be more informative than, for example, a landscape frame” (Li: col. 7, lines 33-51).

13. **Claim 21** is rejected under 35 U.S.C. 103(a) as being unpatentable over Yu et al (US 2002/0126203 A1) with Yasui (US 6,081,615), and further in view of Leow et al (US 7,091,969 B2).

Regarding **claim 21**, Yu with Yasui discloses all elements as mentioned above in claim 1. Yu with Yasui does not disclose using a Voronoi algorithm to determine the shape of the support to be placed on the canvas.

Leow, in the same field of endeavor, teaches using a Voronoi algorithm to determine the shape of the support to be placed on the canvas (see col. 2, lines 19-50; alpha -shape and crust algorithms make use of Voronoi diagrams and Delaunay triangulation to construct triangle mesh. A Voronoi diagram for an arbitrary set of points may be formed from convex polygons created from the perpendicular bisector of lines between nearest neighboring points. Delaunay triangulation forms a mesh using the Voronoi diagrams).

It would have been obvious at the time the invention was made to one of ordinary skill in the art to modify Yu with Yasui to utilize a Voronoi algorithm as taught by Leow, in order for

the output to be topologically correct and convergent to the original surface as the sampling density increases for a "good sample" from a smooth surface (see col. 2, lines 19-50).

14. **Claims 25, 27** are rejected under 35 U.S.C. 103(a) as being unpatentable over Bae et al (US 2002/0126143 A1) in view of Leow et al (US 7,091,969 B2).

Regarding **claim 25**, Bae discloses a computer implemented method implemented within a computer system including memory and CPU for generating a highly condensed visual summary of video regions, comprising:

utilizing the memory and CPU for determining a germ in each of a plurality of images, the germ containing a region of interest (see fig. 5, paragraph [0059]; The synthetic text key frame is generated by summarizing important texts in a video frame in each of the news articles to be inverted into an image; Examiner notes within figure 5, it is noted that the physical text is considered as the germ);

utilizing the memory and CPU for defining a support in each of the video segments, wherein the support is the video segment less the germ;

utilizing the memory and CPU for separating the germ from the video segments (see fig. 5, paragraph [0059]; The synthetic text key frame is generated by summarizing important texts in a video frame in each of the news articles to be inverted into an image);

utilizing the memory and CPU for laying out the germs on a canvas (see fig. 5, paragraph [0059]; the synthetic text key frame is generated by summarizing important texts);

utilizing the memory and CPU for defining a space between the germs (see fig. 5, paragraph [0059]; the synthetic text key frame is generated by summarizing important texts); and

utilizing the memory and CPU for filling in the space of the canvas (see fig. 5, paragraph [0059]; The synthetic text key frame is generated by summarizing important texts in a video frame in each of the news articles to be inverted into an image; Examiner notes that as seen in fig. 5, the germ/region of interest is interpreted to be the text only. Therefore, the corresponding background surrounding the text is considered to be support, since the support is defined as the video segment less the germ. Following this logic, therefore when the summarization of the text occurs, the support along with the germ is separated from the video segments and placed within the synthetic key frame, whereby it fills up the space of the canvas between the germs by having more than one text summarization present in the synthetic key frame. Inherently, if the part of the support and the germ are transferred to the synthetic key frame then naturally, at least one pixel value of the support relative to the closes germ is positioned corresponding to the position of that pixel value relative to the germ) to generate a highly condensed visual summary of the plurality of video segments (see fig. 5, paragraph [0059], synthetic key frame). Bae does not disclose computing boundary curves between the germs using a Voronoi algorithm.

Leow, in the same field of endeavor, teaches computing boundary curves between the germs using a Voronoi algorithm (see col. 2, lines 19-50; alpha -shape and crust algorithms make use of Voronoi diagrams and Delaunay triangulation to construct triangle mesh. A Voronoi diagram for an arbitrary set of points may be formed from convex polygons created from the perpendicular bisector of lines between nearest neighboring points. Delaunay triangulation forms a mesh using the Voronoi diagrams).

It would have been obvious at the time the invention was made to one of ordinary skill in the art to modify the Bae reference to utilize a Voronoi algorithm to compute boundary curves as

taught by Leow, in order for the output to be topologically correct and convergent to the original surface as the sampling density increases for a "good sample" from a smooth surface (see col. 2, lines 19-50).

Regarding **claim 27**, Bac further discloses irregular two dimensional shapes on the canvas (see figs. 5, 6).

Conclusion

15. Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to EDWARD PARK whose telephone number is (571)270-1576. The examiner can normally be reached on M-F 10:30 - 20:00, (EST).

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Samir Ahmed can be reached on (571) 272-7413. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

Edward Park
Examiner
Art Unit 2624

/Edward Park/
Examiner, Art Unit 2624

/Brian Q Le/
Primary Examiner, Art Unit 2624